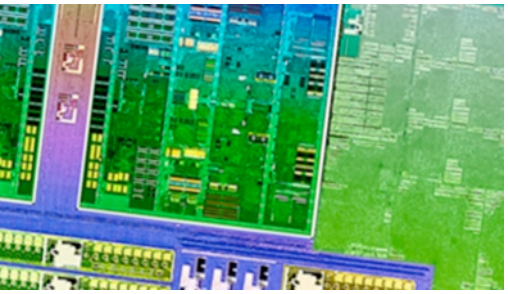
SAND2013-8026P



Sandia National Laboratories

Exceptional service in the national interest

Design Space Exploration with Proxy/Proto Architecture Models and miniApps

James Ang, Ph.D., Senior Manager (acting)
Extreme Scalable Computing Group
Sandia National Laboratories
Albuquerque, NM

ASCR Modeling and Simulation of Exascale Systems and Applications
September 18-19, 2013
University of Washington

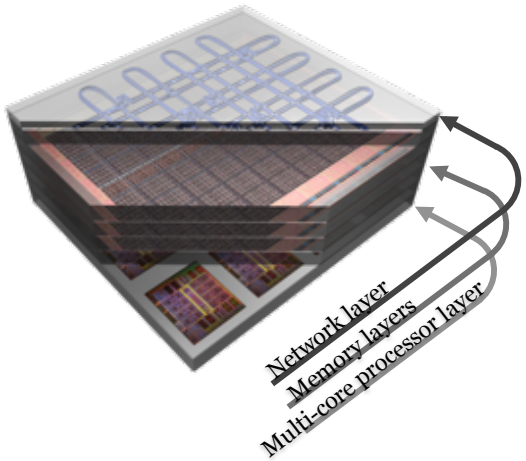
COMPUTER ARCHITECTURE LABORATORY
EXASCALE DESIGN SPACE EXPLORATION

U.S. DEPARTMENT OF ENERGY | Office of Science

National Nuclear Security Administration

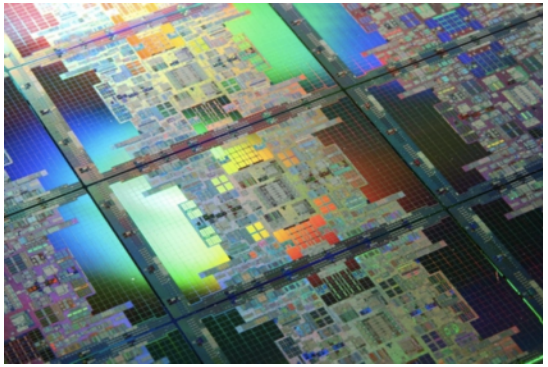
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Exascale Challenges



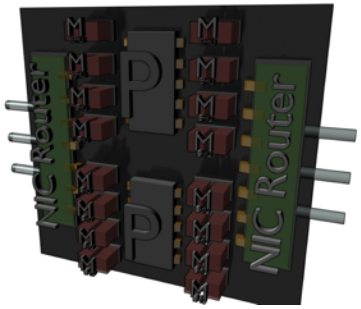
We need to Motivate and *Influence* Architectural Changes

- COTS Processor Architectures
- HPC System Architectures



Our Investments Include Software

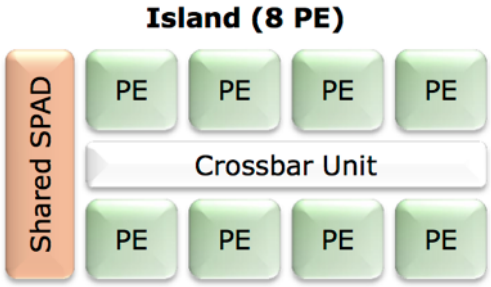
- We cannot just develop new Exascale Architectures and *Throw it over the wall* to our application and system SW developers

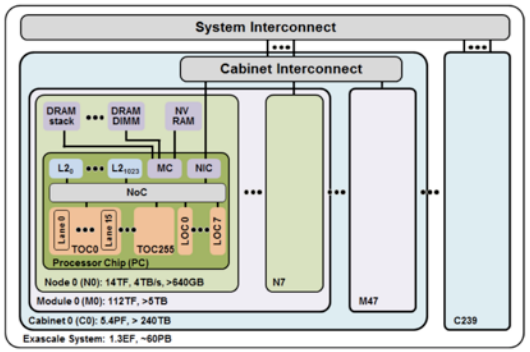


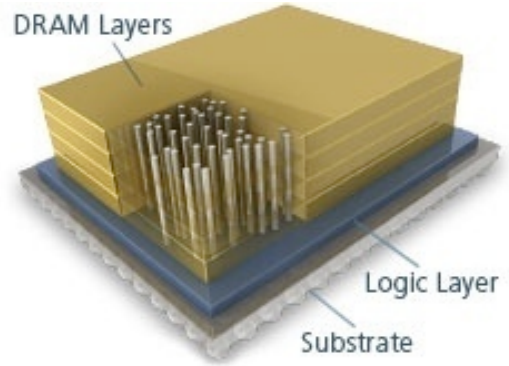
Paths to *Influence* COTS Development

Fund R&D Projects with Industry

- Fast Forward and Design Forward R&D Projects with Industry
  - Improvement on Original ASCI Path Forward Program
  - National Laboratory Staff collaborate with Industry Partners via Co-Design activities, Proxy Applications, Proxy/Proto Architectures, BSM, X-Stack, OS/R, et al
- DOE establishing R&D Projects with Micron
- Explore SoC options for HPC COTS processors
  - Discussions with ARM Holdings and several SoC companies, including some "traditional" companies such as Nvidia and AMD



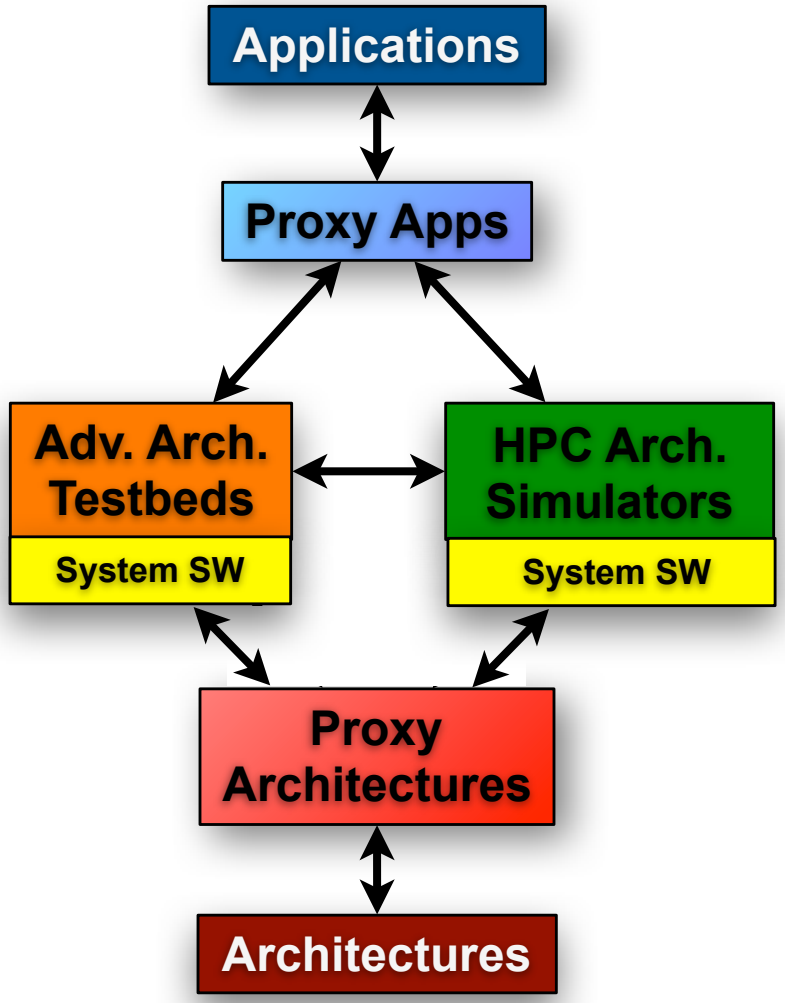




Define the HPC Co-design Methodology

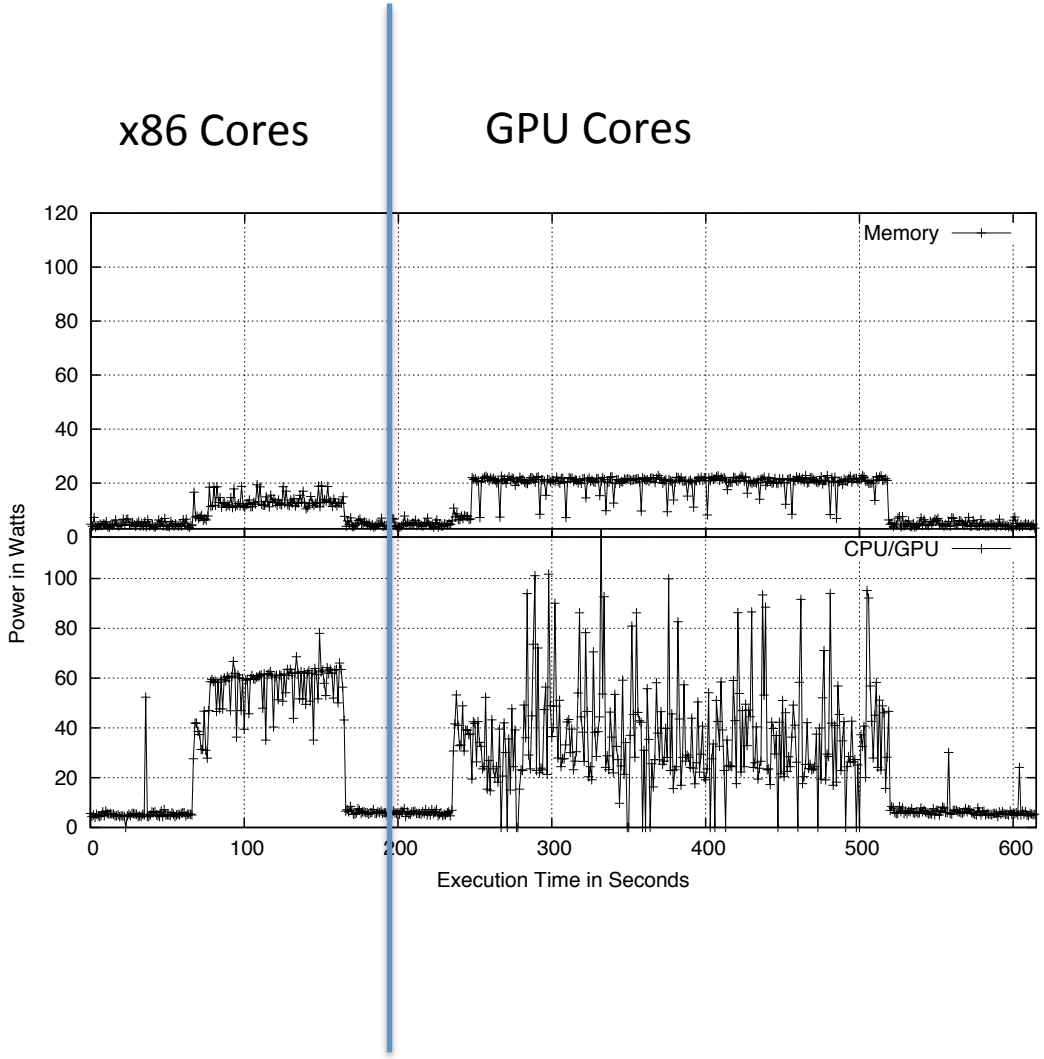
- Key Co-design Capabilities
  - Proxy/Peer mini Applications
  - Development and evolution to represent mission needs
- HPC Architectural Simulators
  - Flexible to accommodate fidelity/ speed tradeoffs
- Proxy/Proto Architectures to explore advanced concepts
  - Abstract machine models
- Advanced architecture testbeds
  - Evolving representation of vendor "state of the art"
  - Support agile system software R&D

X-stack, OS/R, BSM, Execution Model Projects



Mantevo miniFE on AMD *Trinity* APU

- MiniFE is one of Sandia's mini-apps
  - Applications designed to represent core functionality of larger production apps
- First execution only using x86 cores (left part of graph)
- Second execution only using GPU cores (right part of graph)
  - Note, GPU kernel is launched from an x86 core
- Application Energy Profiles
  - Profile of Memory – Top graph
  - Profile of CPU – Bottom graph
- This is an example of the type of application analysis PowerInsight will enable



Computer Architecture Lab: Scope

- Modeling and Simulation Infrastructure:
  Assemble a common set of tools to establish a unified capability *for quantitative analysis of the design space* for hardware and algorithmic/ software design
- Computer Architecture Research and Exploration:
  Lead the definition, development and evolution of *proxy/proto architectures* to facilitate non-proprietary exploration of advanced architecture concepts
- Proving Ground for Industry Technology Options:
  Provide computer architecture staff to function as technical liaisons between the *Architecture Fast Forward and Design Forward* awardees and ESCE RD&EI projects
- Nexus for Co-design/Arch interaction:
  Act as a liaison among the *co-design centers/projects* and commercial *computing technology providers* to help establish interactions

Fast Forward Program

- Objective: *Accelerate transition of innovative ideas from processor and memory architecture research into products*
- Evaluate advanced research concepts and develop quantitative evidence of their benefit for DOE applications (using Co-design Proxy apps)
  - Engage DOE application teams to understand technology trends constraints (how it impacts their code development)
  - Understand how to *program* these new features
  - Quantitative evidence to lower risk to adoption of innovative ideas by product teams
- Processor and Memory Fast Forward Projects
  - AMD – Mike Schulte and Mike Ignatowski, Lab TR – Arun Rodrigues and Jim Ang
  - IBM – Ravi Nair, Jaime Moreno, and Doug Joseph, Lab TR – Bronis de Supinski
  - Intel – Alan Gara and Shekhar Borkar, Lab TR – Sudip Dosanjh and Scott Hemmert
  - NVIDIA - Bill Dally and Steve Keckler, Lab TR John Shalf and Nick Wright

Design Forward Program

Design Forward Technology Focus Areas:

- System design and integration:
  - Overall System Architecture
  - Resilience and Reliability
  - Packaging Density
  - Programming Environment
  - Energy Utilization
  - Data Movement through the System
  - System Software
- Interconnect Networks:
  - Overall Interconnect Architecture
  - Interconnect Integration with Processor and Memory
  - Multiple Communication Library Progression and Interaction
  - Interconnect Fabrics and Management
  - Protocol Support
  - Scalability

CAL *Collaborates* with FastForward (& DF) Partners and ESCE* RD&EI† projects

- The CAL Project has responsibility to develop Proxy/Proto Architecture models that represent FF/DF designs
- The DOE ESCE Co-design Projects have responsibility to develop Proxy Apps that are represent their real apps
- Collaborate with Fast Forward & future Design Forward Projects to understand ability to tune Proto Architecture models to represent their proprietary advanced architecture concepts in our simulators
- Share Proxy/Proto Architectures with ESCE RD&EI projects

* Extreme Scale Computing Effort

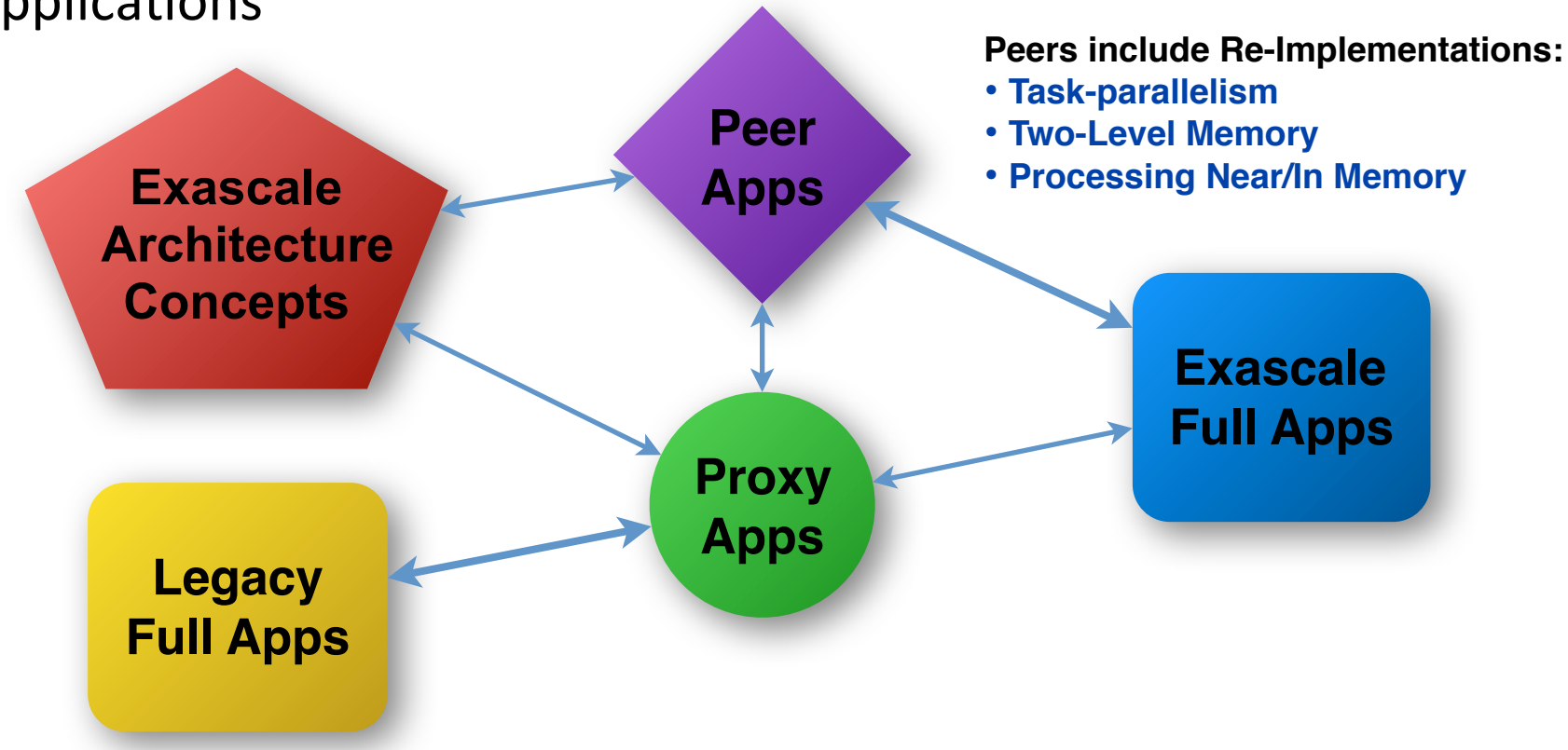
† Research, Development, Integration and Engineering

CAL is also Collaborating with the DOE Co-design Projects

- The three ASCR Co-design Centers and the ASC Co-design Project have responsibility to define and develop Proxy Apps that are representative of their real applications
  - CAL bridges Fast Forward & Design Forward Architecture-centric Projects and App-centric Co-Design Projects
- CESAR – Center for Exascale Simulation of Advanced Reactors
- ExMatEx – Exascale Co-Design Center for Materials in Extreme Environments
- ExaCT: Center for Exascale Simulation of Combustion in Turbulence
- ASC Co-design Project – Working on Tri-Lab coordination within the ASC/IC and ASC CSSE program elements

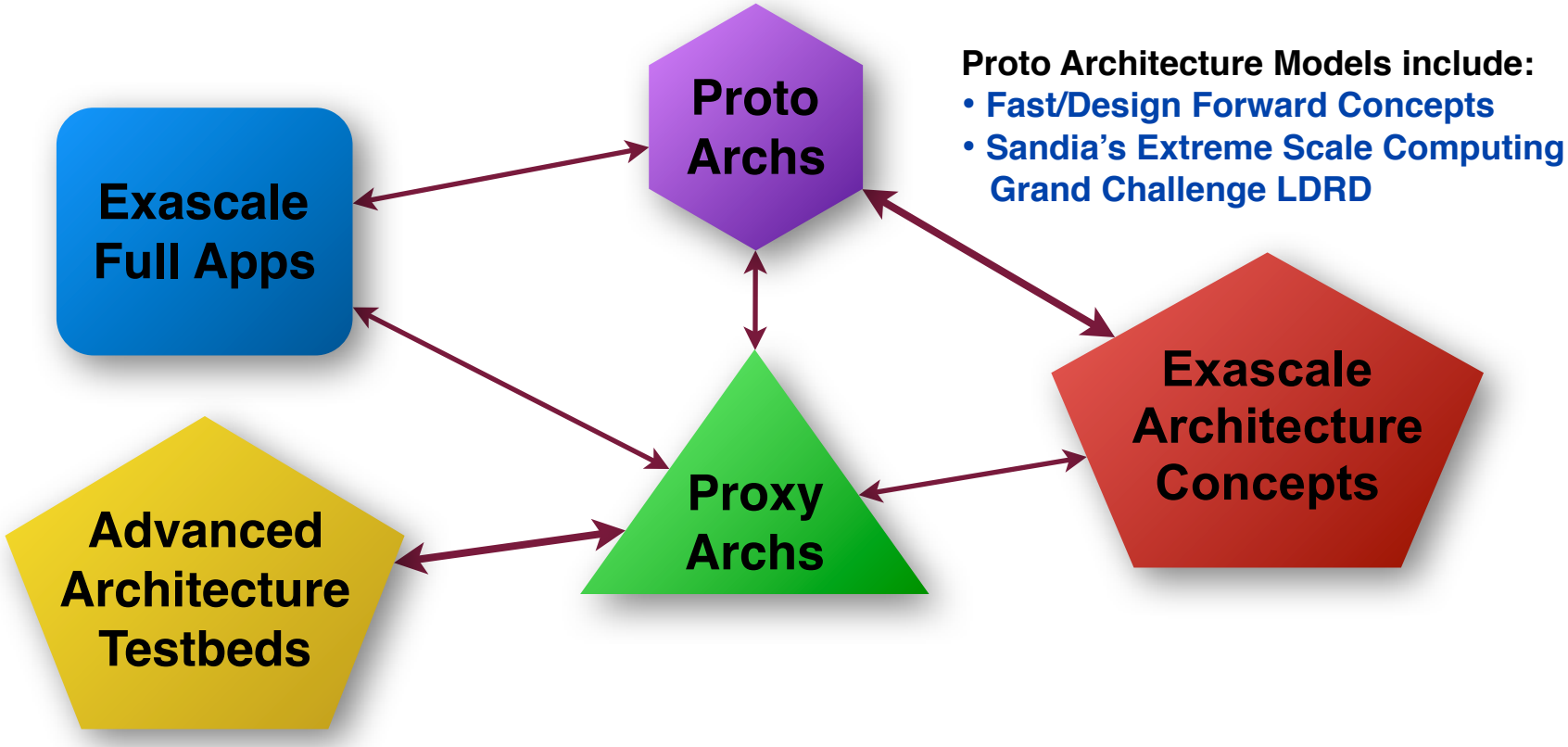
Application Abstractions: MiniApps: Proxy and Peers

- Relationship among Full Applications, Proxy and Peer mini Applications

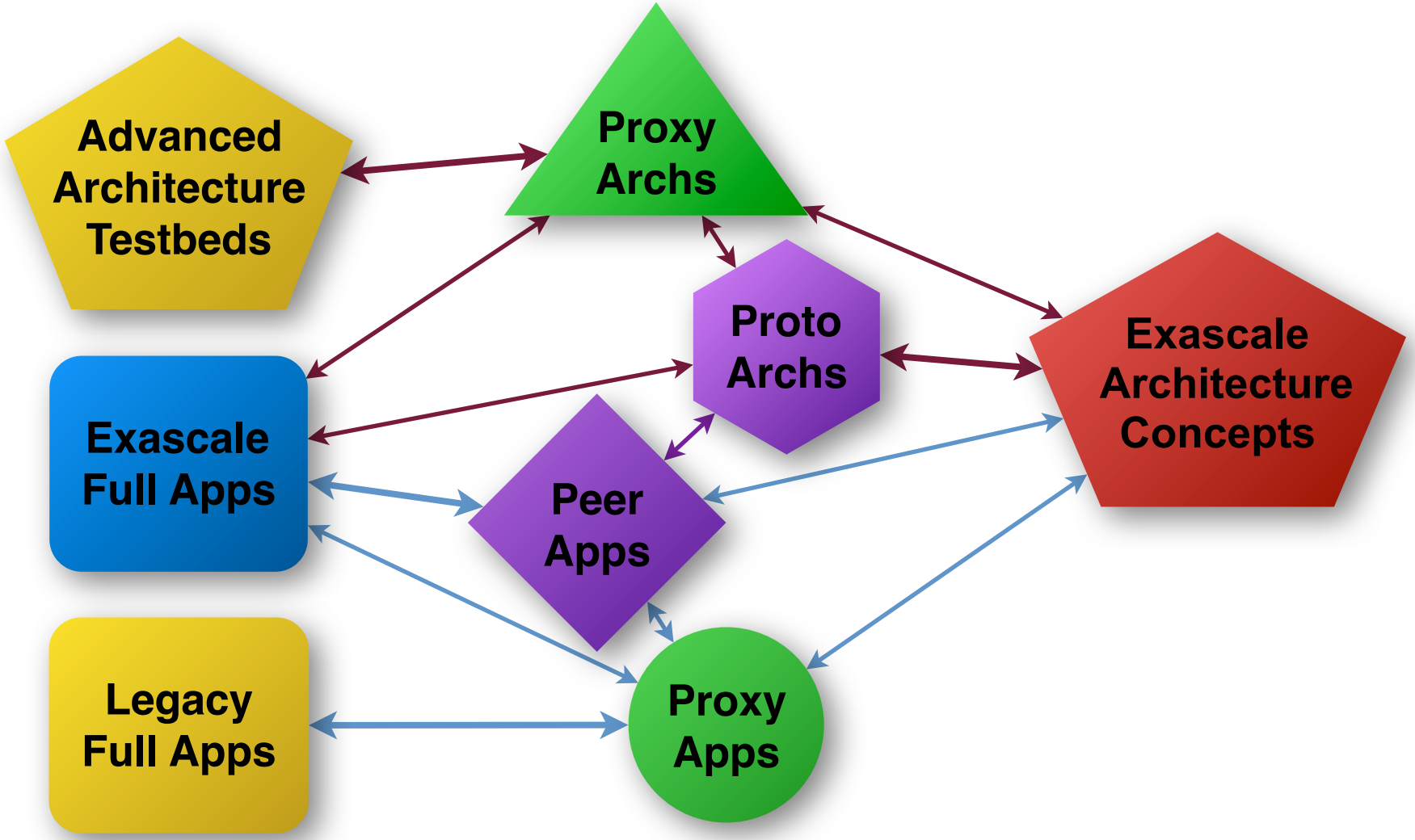


Architecture Abstractions: Proxy and Proto Architectures

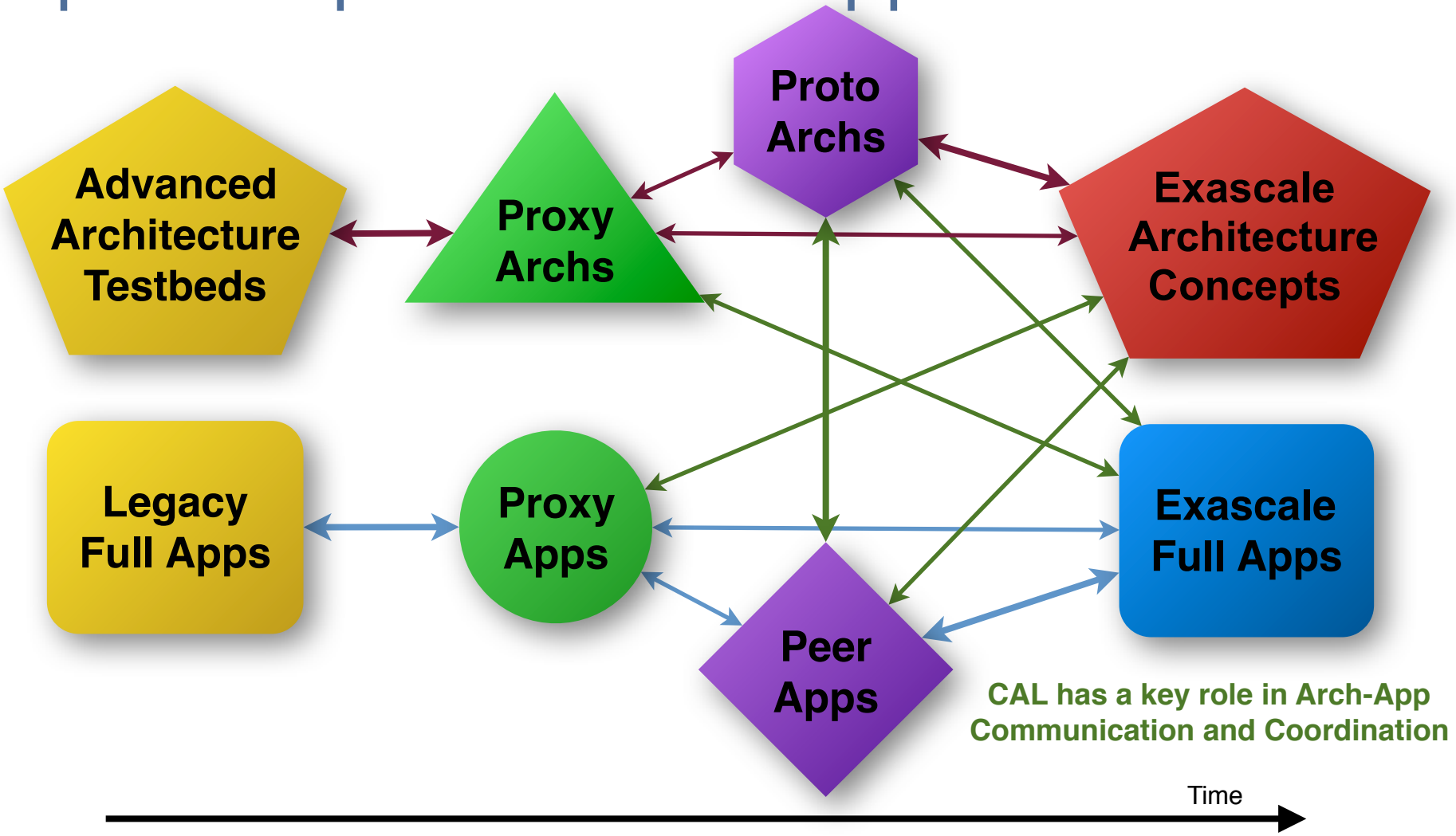
- Relationship among Current, Future, Proxy and Proto Architectures for Design Space Exploration



Overview: Architecture Design Space Exploration and Applications



Another View Architecture Design Space Exploration and Applications



Organizing Committee Questions

- What is the major contribution of your research?
  - Focus on how Co-design can influence future HW Architectures
- What are the gaps you identify in the research coverage in your area?
  - Prototype and Experimental architecture testbeds
  - Proxy miniApps are only a proxy . . .
- What is the bigger picture for your research area? (i.e., identify synergistic projects, complimentary projects in technical sense, etc.)
  - Hardware Dimensions of Co-design apply to both:
    - Processor/Node architecture
    - System Architecture
  - Software Dimensions of Co-design apply to both:
    - Applications
    - System SW

Organizing Committee Questions - cont.

- How do you see cross-pollination across projects funded by different funding agencies?
  - DARPA UHPC/X-Caliber -->
  - SNL/LDRD Extreme Scale Computing Grand Challenge --> ?
- What is the one thing that would make it easier/possible to leverage/use the results of other projects to further your own research?
  - Exercise and Contribute to Co-design capabilities: SST, Mantevo
- What would you like to most see solved/addressed other than what they are working on?
  - Strategy to shift from development of Proxy Applications "Peers" to development of Exascale Applications